

On the spatial correlation between areas of high coseismic slip and aftershock clusters of the Maule earthquake $M_w=8.8$

Javier E. Contreras-Reyes

Servicio Sismológico Nacional
Departamento de Geofísica
Universidad de Chile
Santiago, Chile
Email: jecontrr@mat.puc.cl

Adelchi Azzalini

Dipartimento di Scienze Statistiche
Università di Padova
Padova, Italia
Email: azzalini@stat.unipd.it

Abstract

We study the spatial distribution of clusters associated to the aftershocks of the megathrust Maule earthquake M_W 8.8 of 27 February 2010. We used a recent clustering method which hinges on a nonparametric estimation of the underlying probability density function to detect subsets of points forming clusters associated to high density areas. In addition, we estimate the probability density function using a nonparametric kernel method for each of these clusters. This allow us to identify a set of regions where there is an association between frequency of events and pre-seismic locking. Specifically, our results suggest that high coseismic slip spatially correlates with high aftershock frequency.

Key words: Nonparametric clustering, Kernel density estimation, Maule 2010 Earthquake, aftershock distribution, slip model.

1 Introduction

The most recent large Chilean earthquake occurred on February 27th, 2010 ($M_w = 8.8$) and propagated northward and southward achieving a final rupture length of about 450 km (35–37°S; e.g., Lay et al., 2010; Moreno et al., 2010; Contreras-Reyes and Carrizo, 2011). The earthquake filled a seismic gap (Ruegg et al., 2009) that has experienced little seismic activity since 1835, when it broke with an estimated magnitude of $M_w \sim 8.5$ (Darwin, 1851).

Moreno et al., (2010) show a spatial correlation between coseismic slip and pre-seismic locking of the Maule 2010 event using geodetic data. The authors showed that the two regions of high coseismic slip of the Maule earthquake appeared to be fully locked before the earthquake, and they found a zone that was creeping inter seismically with consistently low coseismic slip bridged by the rupture between those regions. Subsequent geodetic studies have established that the main coseismic slip patch (>15 m) is located in the northern part of the rupture area, with a secondary concentration of slip to the south (5-12 m) (e.g. Lay et al., 2010; Delouis et al., 2010; Lorito et al. 2011; Vigny et al., 2011; Moreno et al., 2012). Lorito et al. (2011) concluded that increased stress on the unbroken southern patch may have increased the probability of another great earthquake there in the near future. In addition, Moreno et al. (2012) suggests that coseismic slip heterogeneity at the scale of single asperities appear to indicate seismic potential for future great earthquakes.

The aim of the present work is to examine the spatial correlation between areas of high coseismic slip and the aftershock frequency (seismic clusters). To achieve this, we use a clustering methodology to identify the distribution of clusters in the rupture area of the 2010 earthquake. The hypocenter data are taken from the SSN (Servicio Sismológico Nacional de Chile), and we use the most recent and accurate coseismic model of Moreno et al. (2012). Clustering techniques represent a long-established area of statistical methodology, with contributions in recent years also from the other disciplines, notably machine learning. Applications of these techniques range over an enormous set of disciplines, both in the natural sciences and in the social sciences. A standard account is the book of Kaufman & Rousseeuw (1990), which is focused mainly on the more classical methods, based on the notion of ‘distance’ between objects.

An alternative, more recent approach is the model-based clustering formulation which regards the observed data as generated by a probability distribution a mixture of multivariate random variables having distribution belonging to some parametric family. More specifically, for any point x in the multivariate set of possible observations, we associate a probability density function $f(x)$ and assume that a decomposition of type

$$f(x) = \sum_{j=1}^J p_j f_j(x)$$

exists, where the p_j ’s are mixing probabilities and $f_j(x)$ denotes the j -th component of the mixture, which also corresponds to the j -th cluster of points, for $j = 1, \dots, J$. In the implementation of this approach, the most common option is to adopt the multivariate Gaussian assumption for each of the components $f_j(x)$, and their parameters are estimated with an EM-type algorithm. From our viewpoint, an especially useful account to this approach is provided by Dasgupta & Raftery (1998), although this kind of formulation is somewhat older. By applying the model-based clustering approach to data on the earthquakes in the coastal area of central California, the authors have obtained six clusters, some of which are clearly linked to the faults described in the literature, such as San Andreas, Calveras and Hayward; however one or two of their clusters do not correspond to some already identified area.

In a further approach to the clustering problem, the notion of an underlying density function $f(x)$ is retained, but the assumption that $f(x)$ is a mixture of components is removed, and so also the connected parametric assumption of the components. Therefore, this approach is based on the construction of a non-parametric estimate $\hat{f}(x)$ of $f(x)$, and the association of a cluster to each of the observed modes of the density. Several alternative estimates of this type exists,

but we shall concentrate on the kernel-type estimate which is the most popular and conceptually very intuitive; an introductory account to the kernel-type density estimation is provided by Bowman & Azzalini (1997). In the univariate case, the kernel estimate is of the form

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \varphi(x - x_i; h), \quad (1)$$

where $\varphi(z; h)$ denotes the normal density function with mean 0 and standard deviation h evaluated at point z . The normal density is adopted for simplicity and it could be replaced by another density symmetric about 0 without much effect on the outcome. A much more important role is played by h which in this context is called ‘smoothing parameter’. In the d -dimensional case, $\varphi(z; h)$ is replaced by a multivariate density with 0 mean vector; the simplest and most popular choice is the product of d such terms, with a different smoothing parameter for each of the d components.

Clusters are then formed by the data points associated to the modes of $\hat{f}(x)$, and the clusters are separated by regions of low density of points. This logic is referred to as nonparametric clustering (NPC). Although the NPC principle had been formulated a long time ago, its actual development is quite recent, due to the demanding computational burden.

Specifically, we adopted the NPC formulation of Azzalini & Torelli (2007), which has the advantage of not requiring on input some subjective choices such as the number of existing clusters. After summarizing this proposal of Azzalini & Torelli (2007) and some follow-up work, we apply this methodology to the SSN catalogue data. We focus on the data of the area between Valparaíso and Tirúa [33-38.5°S], aiming the identification of seismic clusters and its spatial relationship with regions of coseismic slip. Finally, we discuss the possible improvements of the adopted methodology in terms of available geophysical data.

2 Methodology

2.1 Nonparametric clustering

As anticipated earlier, we describe briefly the NPC method proposed by Azzalini & Torelli (2007). This works by assuming that the available set of d -dimensional observations $S = \{x_1, \dots, x_n\}$ represent a set of points drawn from a continuous multivariate random variable having an unknown probability density function $f(x)$, $x \in \mathbb{R}^d$. In the case we are concerned with, $x = (\text{latitude}, \text{longitude})$ denotes a geographical position, hence $d = 2$; the data points x_1, \dots, x_n represent the positions of the observed seismic events.

For any given constant α such that $0 \leq \alpha \leq \max_x f(x)$, consider the high density region defined by

$$R_\alpha = \{x : x \in \mathbb{R}^2, f(x) > \alpha\} \quad (2)$$

which has an associated probability $p_\alpha = \int_{R_\alpha} f(x) dx$. The region R_α is, in general, formed by a number m of connected sets, where $m \in \{1, 2, \dots\}$.

Now we let α move along its range. This causes both m and $p = p(\alpha)$ to move accordingly, and we can regard m as a function of p since p is monotonic with respect to α , write $m(p)$. Note that $m(p)$ is instead not a monotonic function. With the additional conventional settings $m(0) = m(1) = 0$, it can be shown that the total number of increments of the step function $m(p)$ is equal to the number M of modes of $f(x)$, hence to the number of clusters, in the sense defined earlier. As α

varies along its range, and so does $p(\alpha)$, the corresponding connected components of $R(\alpha)$ form a hierarchical tree structure.

Translating the above idea into a working methodology requires some additional specifications and algorithmic work. We sketch here the main step of the procedure. Full details of the method are given by Azzalini & Torelli (1997) and its implementation in the language R is provided by Azzalini *et al.* (2011). This procedure will later be referred to as the ‘pdfCluster method’.

1. A nonparametric estimate $\hat{f}(x)$ of the density is obtained from the observed sample S . Any sort of nonparametric estimate can be employed but in practice both the original paper and pdfCluster adopt a kernel-type estimate. This involves to choose a set of kernel bandwidths, one for each component of x . When the target is estimation of the density itself, it is known that the outcome depends critically on the choice of the bandwidth. However in the present context, where $\hat{f}(x)$ is only an intermediate step toward the final target of clustering, the bandwidth is much less influential and its choice can be varied over a quite wide range without affecting the end result.
2. For any α in the range $0 \leq \alpha \leq \max_i \hat{f}(x_i)$, consider the sample analogue of (2) given by

$$S_\alpha = \{x_i : x_i \in S, \hat{f}(x_i) > \alpha\}. \quad (3)$$

The notion of connected subsets of S_α is introduced via the notion of Delaunay triangulation which is built making use of computational geometry tools. This leads to establish a certain set of line segments joining some of the data points in such a way that, for any two data points, there always exists a path joining them by a sequence of these segments. If x_r and x_s are both points in S_α and the path joining them does not include any point outside S_α , then we say that x_r and x_s are connected, at the given choice of α .

3. The above step is replicated for a grid of values spanning the admissible range of α , from 0 to $\max_i \hat{f}(x_i)$. This generates a mode function $\hat{m}(p)$ and a tree structure of the modes. Notice however that the Delaunay triangulation needs to be determined only once for all α ’s.
4. At the end of the earlier step we have obtained a tree structure of the modes of $\hat{f}(x)$. Moreover, for each of the M modes, we have allocated some elements of the sample S to the given mode. These M subsets of data points form the ‘cluster cores’ to which the remaining points, not belonging to any cluster core, must be aggregated. For each unallocated point x_0 , we must select one of distributions $\hat{f}_1(x), \dots, \hat{f}_M(x)$ which represent the estimated densities of the cluster cores. This allocation is most naturally based on the likelihood ratio, that is we allocate x_0 to the j -th cluster core such that the ratio

$$\frac{\hat{f}_j(x_0)}{\max_{k \neq j} \hat{f}_k(x_0)}$$

is highest. In practice, there are some variant options, depending on whether the $\hat{f}_j(x)$ densities of the cluster cores are updated at each new allocation or not, or according to some intermediate strategy.

The final outcome of the clustering process is represented by the a partition of the sample S into a set of clusters, say C_1, \dots, C_M .

2.2 Density-based silhouette diagnostic

In cluster analysis, the term ‘silhouette’ refers to a diagnostic tool for the validation of the outcome of the clustering process (Rousseeuw, 1987). This technique provides a graphical representation of how well each object lies within the cluster to which it has been allocated; hence it provides an indication of how appropriately the data has been clustered. The idea arises from the comparison of small distance of each observation to the cluster where it has been allocated and a measure of separation from the closest alternative cluster.

Since the original silhouette is based on a distance measure between the observations, it is not adequate for NPC methods where distances play no explicit role. An adaption of the silhouette idea to density-based clustering methods has been proposed by Menardi (2010). The method, called *density-based silhouette* (DBS) diagnostic, is based on the posterior probabilities

$$p_j(x_i) = \frac{\hat{\pi}_j \hat{f}_j(x_i)}{\sum_{j=1}^M \hat{\pi}_j \hat{f}_j(x_i)}, \quad j = 1, \dots, M,$$

where $\hat{\pi}_j$ plays the role of prior probability of cluster C_j ; in practice it is taken to be the proportion of points allocated to C_j . The DBS index of observation x_i is

$$\text{dbs}(x_i) = \frac{\log\left(\frac{p_{j_0}(x_i)}{p_{j_1}(x_i)}\right)}{\max_{k=1, \dots, n} \left| \log\left(\frac{p_{j_0}(x_k)}{p_{j_1}(x_k)}\right) \right|}$$

where j_0 denotes the cluster to which x_i has been allocated, and j_1 refers to the alternative cluster index for which p_j is maximum, $j \neq j_0$.

In our case, after partitioning the SSN data described earlier with the pdfCluster method, we applied the DBS diagnostic to assess the quality of the outcome.

2.3 Temporal analysis

The clustering estimation can be modified depending of the number of days after the mega earthquake and the observations involved in each day may be produce unequal results. Hence, for $t > 1$, we compare the results at day t respect to day $t - 1$ to compare two alternative partitions of the same set. In each comparison it is necessary to keep the same number of clusters at day t and $t + 1$.

We briefly describe four indexes to illustrate the consistency of the NPC method across the time. These indexes perform accuracy of the observed in predicting the correct category, relative to that of random chance. In Table 1, $n(F_i, O_j)$ denotes the number of forecasts in category i that had observations in category j , $N(F_i)$ denotes the total number of forecasts in category i , $N(O_j)$ denotes the total number of observations in category j , N is the total number of forecasts and $i, j = 1, \dots, 5$ are the indexes of the five identified clusters.

$$\begin{aligned} NSS &= \frac{1}{N} \sum_{i=1}^5 N(F_i, O_i), \\ HSS &= \frac{NSS - \frac{1}{N^2} \sum_{i=1}^5 N(F_i)N(O_i)}{1 - \frac{1}{N^2} \sum_{i=1}^5 N(F_i)N(O_i)}, \\ HK &= \frac{N^2 - \sum_{i=1}^5 N(F_i)N(O_i)}{N^2 - \sum_{i=1}^5 N(O_i)^2} HSS \end{aligned}$$

where NSS corresponds to normal skill score with range $[0, 1]$, such that 0 indicates no skill, and 1 indicates perfect score. The index HSS correspond to Heidke skill

Table 1: Multi-category Contingence table with five clusters.

		Observed Category					
		1	2	3	4	5	Total
Forecast Category	1	$n(F_1, O_1)$	$n(F_1, O_2)$	$n(F_1, O_3)$	$n(F_1, O_4)$	$n(F_1, O_5)$	$N(F_1)$
	2	$n(F_2, O_1)$	$n(F_2, O_2)$	$n(F_2, O_3)$	$n(F_2, O_4)$	$n(F_2, O_5)$	$N(F_2)$
	3	$n(F_3, O_1)$	$n(F_3, O_2)$	$n(F_3, O_3)$	$n(F_3, O_4)$	$n(F_3, O_5)$	$N(F_3)$
	4	$n(F_4, O_1)$	$n(F_4, O_2)$	$n(F_4, O_3)$	$n(F_4, O_4)$	$n(F_4, O_5)$	$N(F_4)$
	5	$n(F_5, O_1)$	$n(F_5, O_2)$	$n(F_5, O_3)$	$n(F_5, O_4)$	$n(F_5, O_5)$	$N(F_5)$
Total		$N(O_1)$	$N(O_2)$	$N(O_3)$	$N(O_4)$	$N(O_5)$	N

score (Brier & Allen, 1952) with range $(-\infty, 1]$, 0 indicates no skill, and 1 indicates perfect score. Measures the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance. This is one form of a generalized skill score, where the score in the numerator is the number of correct forecasts, and the reference forecast in this case is random chance.

The index HK correspond to Hanssen & Kuipers discriminant (Hanssen & Kuipers, 1965) with range $[-1, 1]$, such that 0 indicates no skill, and 1 indicates perfect score. Similar to the Heidke skill score, except that in the denominator the fraction of correct forecasts due to random chance is for an unbiased forecast. Hubert & Arabie (1985) noticed that the Rand index is not corrected for chance that is equal to zero for random partitions having the same number of objects in each class. They introduced the corrected Rand index, whose expectation is equal to zero. The adjusted Rand index is based on three values: the number r of common joined pairs in O and F, the expected value $E(r)$ and the maximum value $\max(r)$ of this index, among the partitions of the same type as O and F. It leads to the formula

$$HA(O, F) = \frac{r - E(r)}{\max(r) - E(r)}$$

where

$$r = \sum_{i=1}^5 \sum_{j=1}^5 \frac{n_{i,j}(n_{i,j} - 1)}{2}, \quad E(r) = \frac{1}{2} \frac{|N(O)| \times |N(F)|}{n(n-1)},$$

$$|N(O)| = \sum_{i=1}^5 \frac{n_i(n_i - 1)}{2}, \quad |N(F)| = \sum_{j=1}^5 \frac{n_j(n_j - 1)}{2}$$

and $\max(r) = \frac{1}{2}(|N(O)| + |N(F)|)$. This maximum value is questionable since the number of common joined pairs is necessarily bounded by $\inf\{|N(O)|, |N(F)|\}$, but $\max(r)$ insures that the maximum value of HA is 1 when the two partitions are identical.

3 Results

Our numerical work is based on data extracted from the SSN catalogue, available at <http://ssn.dgf.uchile.cl/>. Specifically, we have considered 6,714 after-shocks in a map $[32-40^\circ\text{S}] \times [69-75.5^\circ\text{E}]$, for a period between 27 February 2010 and 13 July 2011 (see Fig. 4) and for local magnitudes $M_l \geq 2.0$. All of these observations have been previous processed by the SSN with SEISAN 8.3 software considering the information provided by 22 stations located in a map with coordinates $[-33.32, -39.80]$ latitude and $[-70.29, -73.24]$ longitude.

For the numerical processing, we used the R computing environment software (R Development Core Team, 2011). Most of the work was done using the R

package `pdfCluster` by Azzalini et al. (2011); this package comprises the functions `pdfCluster` for the NPC method, `dbs` for DBS diagnostics and `kepdf` for kernel density estimation. In addition, R provides `kruskal.test` function for a nonparametric-type analysis of variance and `wilcox.test` function for individual test of means.

3.1 Clustering process

The top panel of Figure 4 displays the geographical map of the area of interest with the points denoting the locations of the events. These points have been clustered using the `pdfCluster` method described in Section 2, leading to the different coloring. The bottom portion of the figure shows cluster tree and the silhouette diagnostic. These indicate a lack of clear separation among clusters, which is not surprising, given the close geographical proximity of the clouds of points visible in the top panel of the figure.

In the second stage of our numerical work, we have introduced two variants. One was to consider only the areas where slip did take place in a non-negligible form, since our aim was exactly to examine the implications of the slip model. In addition, some numerical exploration has indicated that log-transformation of the quantities, slip and density function, lead to a more meaningful outcome. Since for a number of points the slip value is 0, we adopted the modification commonly used in this case of adding a small positive quantity, that is working $\log(k + \text{slip})$, where k is some small quantity. Since slip is measured in metres, then we adopted $k = 0.01$ which represents a perturbation of only 1 cm of the original data. In the following, the term log-slip will be used for referring to $\log(0.01 + \text{slip})$.

Furthermore, association between slip and events density can be examined in two different ways. One is to choose a regular grid of points in the region of interest, and evaluate these variables or their log-transforms over this grid. The other option is to evaluate these variables at the observed points of the seismic events.

Figure 5 refers to the first form, for three choices of the geographical area over which the grid of points is constructed. More precisely, the sets of points for which the computations have been performed have been obtained as the intersection of rectangular grids of sizes (192×214) , (149×214) and (119×214) with the three regions shown on the left side of Figure 5, of different geographical size. This process led then to consider three non-rectangular grids, comprising 19630, 10696 and 4812 points, respectively. The area covered by first grid includes the largest number of points associated of the seismic events of the data set, while the last one refers to the area with highest concentration of events. The right side of the figure displays the scatter plots of log-slip and log-density of the points, separately for each cluster. Only clusters labelled No. 1, 2, 3, 6, and 8 are considered here, since the other clusters of Figure 4 have been dropped, for the reason explained earlier.

Even if the selection of the three grids is somewhat subjective, the overall indication provided from Figure 5 provides convincing evidence of the presence association between the variables under consideration, specifically clusters labeled with No. 1, 2, and 6. This association is becoming more and more marked as we move down from the first to the last grid, that is when we focus on the area with greater intensity of events. The type of association is definitely non-linear, and so admittedly it does not lend itself to simple interpretation, but it is clearly present, especially so in the bottom portion of the figure.

Figure 3 refers instead to the second form of comparison, where evaluation of log-slip and log-density is performed at the observed location of events instead of a regular grid of points. Also this figure exhibits some noteworthy features. One is that in the red, black, violet, and grey clusters there exists a clear positive

Table 2: Summary statistics of the slip variable by cluster

Cluster	Mean	Min	Max	S.D.	N
red	6.200	0	16.566	4.329	4165
black	7.579	0.04	13.728	2.623	950
green	0.084	0	1.945	0.331	265
violet	7.572	0	15.043	5.752	149
gray	4.043	0	11.875	3.028	308

Table 3: *P*-values for Wilcoxon test for clusters slip.

	red	black	green	violet	gray
red	-	0	0	0.0181	0
black	0	-	0	0.3497	0
green	0	0	-	0	0
violet	0.0181	0.3497	0	-	0
gray	0	0	0	0	-

association between log-pdf and log-slip. Hence, the maximum slip is associated with high frequency of events. The green cluster does not display any association, presumably so because several events matching with null slip zone where probably have not been involved with the main earthquake; however the slip produced in this zone is lower. Pichilemu city (34.38°S , 72.02°W) is located in the middle of the red cluster, approximately, which where the maximum slip is 16.6 meters. The sky-blue cluster correspond to seismic activity produced by Puyehue volcano eruption (June, 2011). Mocha Island (38.39°S , 73.87°W) is located in the bottom of the gray cluster, which where the maximum slip is 11.9 meters (see Figure 2). In the black and gray clusters, we can see a positive association: the values of log-slip increments in the measure that the log-pdf increase.

3.2 Nonparametric analysis of variance

A positive diagnostic of NPC allows us to implement the Kruskal-Wallis one-way analysis of variance method (KW; Kruskal & Wallis, 1952) by ranks that show the significance existence of slip between the clusters and consider a confidence level to testing what is the high slip of a cluster over the others. The KW is a classical method for testing whether samples originate from the same distribution where the null hypothesis is that the groups from which the samples originate, have the same median. Since it is a non-parametric method, the KW test does not assume a normal population or another distribution, but its purpose is analogous to the one of the classical analysis of variance for normal populations. The KW method consider a test statistic corrected by ties to compute the p-value and, large values of this test statistic produce the reject of the null hypothesis that the median of groups are equal. We made use of this method to test whether the clusters produced by the NPC method are associated to different slip measurements. The numerical outcome of the R function `kruskal.test` was 1861.5 with an associated *p*-value (below 10^{-15}), which provides an extreme indication of heterogeneity of slip among the clusters.

To examine where the differences among the groups are, we make use of the Wilcoxon test (Wilcoxon, 1945). This test perform individual test between two

Table 4: Geodetic distance of clusters from the trench.

Cluster	min	max	mean
red	18.34	326.78	100.67
black	1.89	173.39	74.53
green	98.82	219.64	147.55
violet	1.31	168.93	27.56
gray	3.42	212.91	68.75

groups assuming for a null hypothesis that not exist differences between the two medians. The results are shown in Table 3. If each p -value is consider isolatedly, there is only on non-significant comparison at 5% level, but we must make an allowance for repeated testing; in this case, 10 testing procedures have been performed. The more classical form of allowance for repeated testing is via the Bonferroni correction, which here leads to consider the $0.05/10 = 0.005$ significance level. Therefore, also the value 0.0181 must be regarded as non-significant.

We can see in Table 2 that the red cluster representing the Pichilemu zone has the higher maximum of slip in relation with red (Constitución zone) and violet (Pichilemu’s offshore coast zone) clusters. With respect the gray (Arauco zone) and green (Rancagua city zone) clusters, present the lower values of slip and practically, the green cluster does not present slip.

Figure 1 shows the consistency of NP method along 200 days since the moment of the great earthquake with values higher than 0.89 for HA case, 0.95 for NSS and HSS cases, and 0.7 for HK case. In the first 200 days, some compressions between the clusters estimation at day t versus day $t - 1$ produce lower values of the indexes by the incorporation of one or more new groups related to the added observations.

The pronounced crustal aftershock activity with mainly normal faulting mechanisms is found in the Pichilemu region (Farías et al., 2011; Lange et al., 2012). Lange et al., (2012) consider the processed events between 15 March and 30 September 2010 to estimate local magnitudes (M_l) in the Pichilemu region, where those magnitudes are comparable with the SSN magnitudes for large events. Specifically, a crustal aftershock activity is found in the region of Pichilemu ($\sim 34.5^\circ\text{S}$) where the crustal events occur in a ~ 30 km wide region with sharp inclined boundaries and oriented oblique to the trench. This region, detected after the $M_w = 8.8$ thrust subduction Maule earthquake by Farías et al. (2011), coincide with the coseismic slip models such as Lay et al. (2010), Delouis et al. (2010) and Lorito et al. (2011). Specifically, coincide with one of the maximum patches of slip ($\sim 12\text{m}$, Moreno et al., 2010, 2012). On another hand, the aftershock seismicity parallel to the trench is apparent at 50-120 km distance perpendicular to the trench (see Table 4). Near $\sim 35^\circ\text{S}$ and in the southern part of the rupture at $\sim 38^\circ\text{S}$ occurs significant aftershock activity after the megathrust earthquake. This seismicity occurs in regions of high coseismic slip (see Table 2). Aftershocks and coseismic slip of the Maule 2012 earthquake terminate ~ 50 km south of the prolongation of the subducting Mocha Fracture zone around $\sim (73.5^\circ\text{W}, 38.5^\circ\text{S})$, near of the bottom of gray cluster (see Figure 2).

4 Conclusions

We have proposed an alternative way to clustering the aftershocks seismicity of the 2010 Maule earthquake M_W 8.8. The nonparametric clustering has shown to

be consistent in the measure that the dairy aftershocks events are added in the analysis and we present the diagnostic tools to illustrate this feature. We using a nonparametric kernel method to fit the high empirical aftershock frequency, which were highly correlated with the used coseismic slip model. Our findings can be explored further by considering an extended data set, including the events with delayed effect, and modeling the relationship of high coseismic slip areas and aftershock clusters. This analysis should be considered, in the attempt to help identification of an increasing risk of occurrence of another great earthquake.

Acknowledgments

We like to thank Giovanna Menardi for useful comments and suggestions in connection with the pdfCluster methodology; to Eduardo Contreras-Reyes and Marcos Moreno for useful comments, suggestions and computational help; and Hector Massone for seismicity data base used in this article, respectively.

References

- Azzalini, A., Torelli, N., 2007. Clustering via nonparametric density estimation. *Stat. Comput.* 17, 71-80.
- Azzalini, A., Menardi, G., Rosolin, T., 2011. R package *pdfCluster*: Cluster analysis via nonparametric density estimation (version 0.1-13). URL <http://cran.r-project.org/package=pdfCluster>
- Bowman, A.W., Azzalini, A., 1997. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford University Press, Oxford.
- Brier, G.W., Allen, R.A., 1952. Verification of weather forecast. *Compendium of Meteorology*, Boston, Amer. Meteor. Soc., 841-848.
- Contreras-Reyes, E., Carrizo, A., 2011. Control of high oceanic features and subduction channel on earthquake ruptures along the Chile-Perú subduction zone. *Physics of the Earth and Planetary Interiors* 186, 49-58.
- Darwin, C., 1851. Geological Observations on Coral Reefs, Volcanic Islands and on South America. *Londres, Smith, Elder and Co*, p. 768.
- Dasgupta, A., Raftery, A.E., 1998. Detecting features in spatial point processes with cluster via model-based clustering. *Journal of the American Statistical Association* 93, 294-302.
- Delouis, B., Nocquet, J.M., Vallee, M., 2010. Slip distribution of the February 27, 2010 Mw = 8.8 Maule Earthquake, central Chile, from static and high-rate GPS, InSAR, and broadband teleseismic data, *Geophys. Res. Lett.* 37, L17305, doi:10.1029/2010GL043899.
- Farías, M., Comte, D., Roecker, S., Carrizo, D., Pardo, M., 2011. Crustal extensional faulting triggered by the 2010 Chilean earthquake: The Pichilemu Seismic Sequence. *Tectonics* 30, TC6010, doi:10.1029/2011TC002888.
- Hansen, A.W., Kuipers, W.H.A., 1965. On the relationship between the frequency of rain and various meteorological parameters. Koninklijke Nederlands Meteorologisch Institut, *Meded. Verhand.*

- Hubert, L., Arabie, P., 1985. Comparing Partitions. *Journal of the Classification* 2, 193-218.
- Kaufman, L., Rousseeuw, P.J., 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley.
- Kruskal, W., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47 260, 583-621.
- Lange, D., Tilmann, F., Barrientos, S., Contreras-Reyes, E., Methe, P., Moreno, M., Heit, B., Bernard, P., Vilotte, J., Beck, S., 2012. Aftershock seismicity of the 27 February 2010 Mw 8.8 Maule earthquake rupture zone. *Earth and Planetary Science Letters*, doi:10.1016/j.epsl.2011.11.034.
- Lay, T., Ammon, C.J., Kanamori, H., Koper, K.D., Sufri, O., Hutko, A.R., 2010. Teleseismic inversion for rupture process of the 27 February 2010 Chile (Mw= 8.8) earthquake. *Geophys. Res. Lett.* 37, L13301, doi:10.1029/2010GL043379.
- Lorito, S., Romano, F., Atzori, S., Tong, X., Avallone, A., McCloskey, J., Cocco, M., Boschi, E., Piatanesi, A., 2011. Limited overlap between the seismic gap and coseismic slip of the great 2010 Chile earthquake. *Nat. Geosci.* 4, 173-177. doi:10.1038/ngeo1073.
- Menardi, G., 2010. Density-based silhouette diagnostics for clustering methods. *Stat. Comput.* 21, 295-308.
- Moreno, M., Rosenau, M., Oncken, O., 2010. Maule earthquake slip correlates with pre-seismic locking of Andean subduction zone. *Nature* 467, 198-202. doi:10.1038/nature09349.
- Moreno, M., Melnick, D., Rosenau, M., Baez, J., Klotz, J., Oncken, O., Tassara, A., Chen, J., Bataille, K., Bevis, M., Socquet, A., Bolte, J., Vigny, C., Brooks, B., Ryder, I., Grund, V., Smalley, B., Carrizo, D., Bartsch, M., Hase, H., 2012. Toward understanding tectonic control on the Mw 8.8 2010 Maule Chile earthquake. *Earth and Planetary Science Letters*, doi:10.1016/j.epsl.2012.01.006.
- Pollitz, F.F., Brooks, B., Tong, X., Bevis, M.G., Foster, J.H., Bürgmann, R., Smalley, R.J., Vigny, C., Socquet, A., Ruegg, J.-C., Campos, J., Barrientos, S., Parra, H., Baez Soto, J.-C., Cimbaro, S., Blanco, M., 2011. Coseismic slip distribution of the February 27, 2010 Mw 8.8 Maule, Chile earthquake. *Geophys. Res. Lett.* 38, L09309, doi:10.1029/2011GL047065.
- R Development Core Team, 2011. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53-65. doi:10.1016/0377-0427(87)90125-7.
- Ruegg, J., Rudloff, A., Vigny, C., Madariaga, R., de Chabali er, J.B., Campos, J., Kausel, E., Barrientos, S., Dimitrov, D., 2009. Interseismic strain accumulation measured by GPS in the seismic gap between Constituci n and Concepci n in Chile. *Physics of the Earth and Planetary Interiors*

Vigny, C., Socquet, A., Peyrat, S., Ruegg, J.-C., Métois, M., Madariaga, R., Morvan, S., Lancieri, M., Lacassin, R., Campos, J., Carrizo, D., Bejar-Pizarro, M., Barrientos, S., Armijo, R., Aranda, C., Valderas-Bermejo, M.-C., Ortega, I., Bondoux, F., Baize, S., Lyon-Caen, H., Pavez, A., Vilotte, J.P., Bevis, M., Brooks, B., Smalley, R., Parra, H., Baez, J.-C., Blanco, M., Cimbaro, S., Kendrick, E., 2011. The 2010 Mw 8.8 Maule Mega-Thrust Earthquake of Central Chile, Monitored by GPS. *Science* 332 (6036), 1417-1421. doi:10.1126/science.1204132.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics* 1, 80-83.

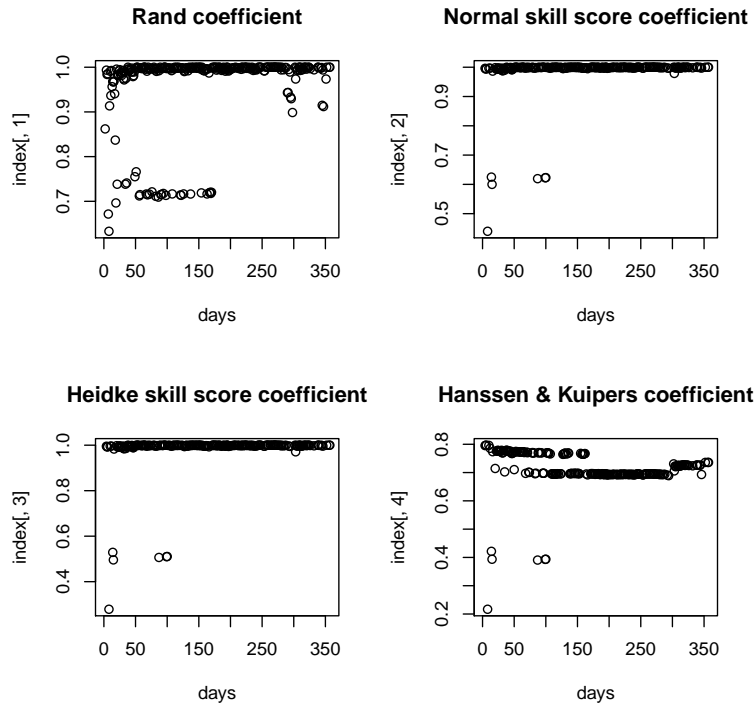


Figure 1: Plots of indexes of NPC results comparisons of data set at t day versus $t - 1$ day.

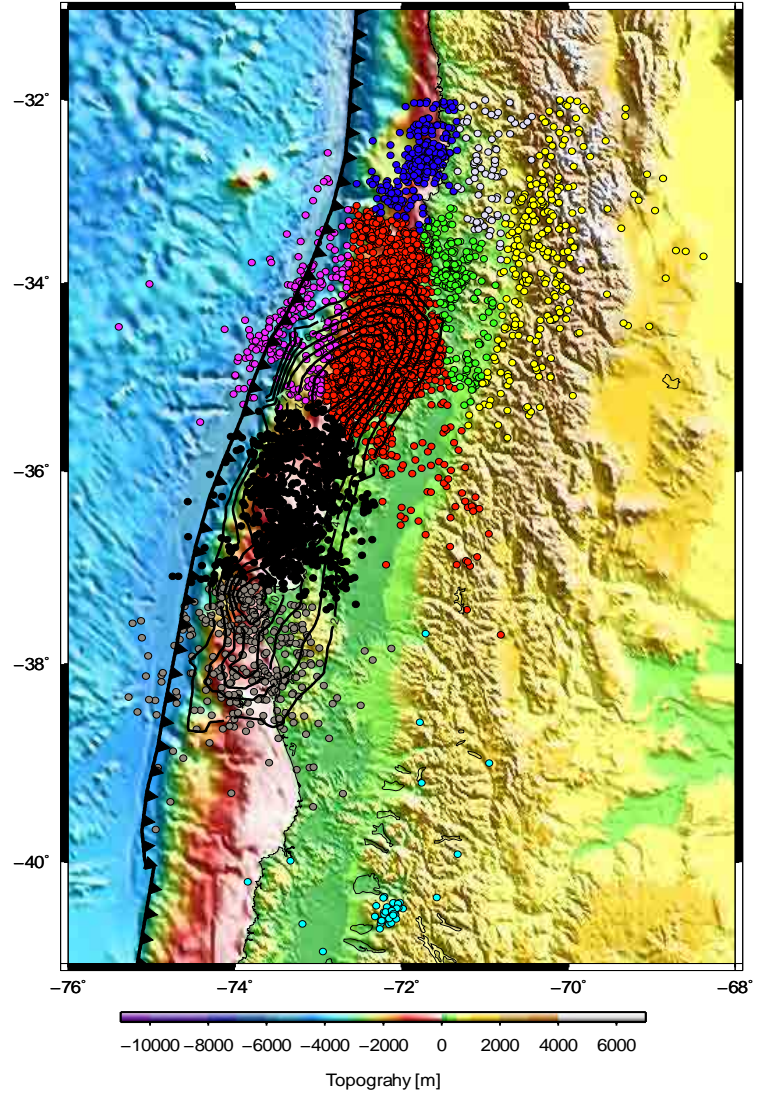


Figure 2: Map of the Chile region analyzed for slip and post-seismicity correlation with clustering events. The black-triangle line correspond to the trench.

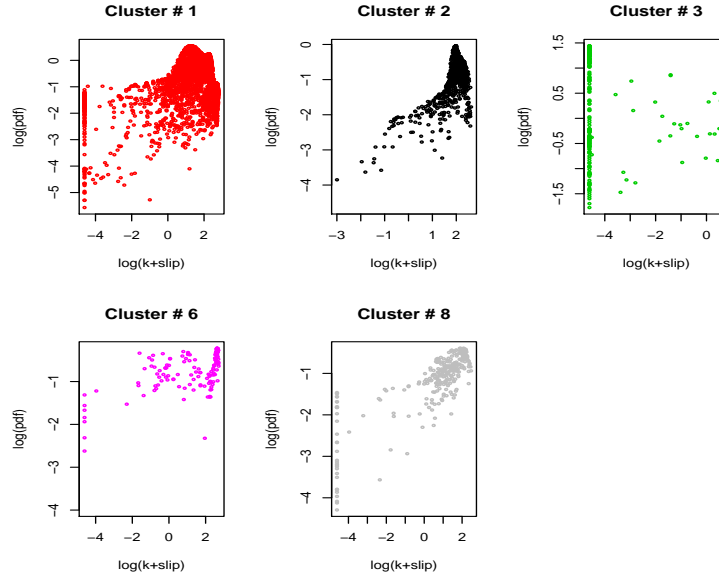


Figure 3: Relationship between log-slip and post-seismicity log-frequency for the clusters obtained for the NPC method of the first plot of Figure 4 that consider the slip model. The points with slip = 0 produce vertical strips of points at abscissa $\log(0.01) \approx -4.6$

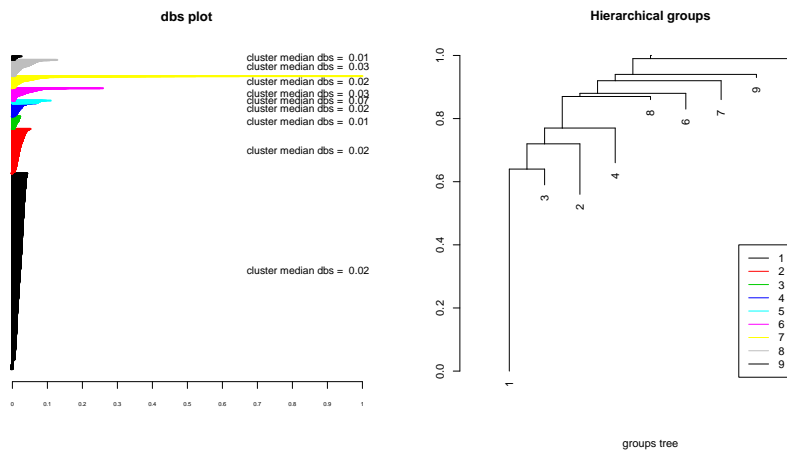


Figure 4: Plots of NPC method results.

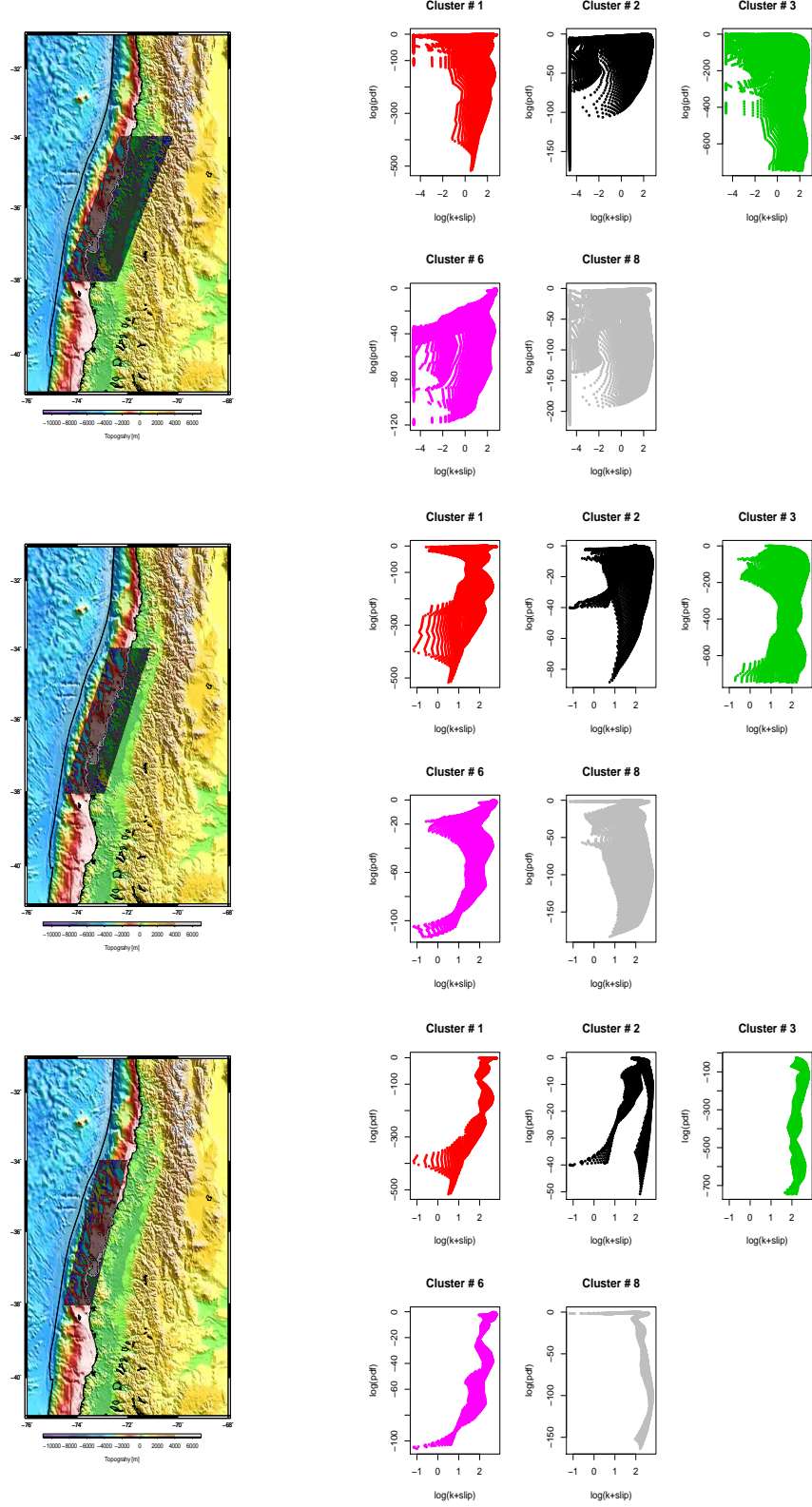


Figure 5: Left: Three grids considered for the NPC method (blacked shadows). Right: Relationship between \log -slip and post-seismicity \log -frequency for each grid. The points with $\text{slip} = 0$ produce verticals strips of points at abscissa $\log(0.01) \approx -4.6$ in some of the top plots